

Experimental Design : Optimizing Quantities of Interest to Reliably Reduce the Uncertainty in Model Input Parameters

Scott Walsh
Advisor: Troy Butler

Contents

1	Introduction	2
2	A Measure-Theoretic Stochastic Inverse Problem	3
2.1	A Mathematical Formulation	3
2.2	Numerical Approximation of Solutions	7
3	Precision and Accuracy	9
3.1	μ_Λ -measure of support	11
3.2	Skewness	13
4	Multicriteria Optimization	15
4.1	Discrete Optimization	16
4.2	Numerical Examples	16
4.2.1	Linear Example	16
4.2.2	Nonlinear	17
5	Maps Defined by the Solution of a Physics-Based Model	20
5.1	Time Dependent Diffusion	20
5.1.1	The Model	20
5.1.2	The Inverse Problem	21
5.1.3	The Prediction Problem	24
5.2	Coastal Ocean Model	25
5.2.1	The Model	25
5.2.2	The Inverse Problem	26
6	Conclusion	29

1 Introduction

With an abundance of computing resources available today mathematical models are capable of simulating extraordinarily complex physical systems. These complex models in turn induce a set of model input parameters, many of which are subject to uncertainty. For example, the diffusion coefficient of a model simulating the heating of a thin plate by some external source is subject to uncertainty, possibly due to the manufacturing imperfections of the thin plate. The Manning's n parameters of a coastal storm surge model may initially be determined by extensive exploration of the coastal boundary, however each passing storm has the potential to significantly alter regions of this coastal boundary introducing uncertainty into these parameters. Our goal is to reduce the uncertainty in these model input parameters, therefore improving the predictive capabilities of the model.

One approach to reducing the uncertainty in these model input parameters is to gather quantities of interest (QoI) of the physical system and use this data to inform us about the parameters that may have produced this data. This data may be a temperature measurement at a point in space time or a maximum water elevation at a given location in the physical domain. As our model input parameters are uncertain, so is the data we gather. The gathered data is uncertain for a multitude of reasons, e.g., measurement instruments have finite precision, instrument locations in space time are subject to uncertainty. This uncertainty in each QoI gathered may also be a function of the region of the parameter space that produced this physical reality. Different observed solutions (corresponding to different regions of model input parameters) may produce physical processes that are not being modeled that pollute measurements even more. For example, some regions of a parameter space of a coastal storm surge model may produce large and chaotic waves that pollute measurements.

This uncertainty in the QoI produces a set valued inverse problem. In [2, 5] the foundation for a measure theoretic approach to solving this set valued inverse problem is carefully developed. With this foundation built on well understood mathematical concepts, the uniqueness of such set valued inverse solutions is proved, under suitable conditions. However, each solution is dependent on the data used in the inverse problem, i.e., different sets of data produce different inverse solutions whose geometric properties may vary substantially. In this paper we develop an approach for determining an optimal choice of QoI to produce a *precise* and *accurate* inverse solution, i.e., reduce the uncertainty in our model parameters.

In Section 2 we establish the foundation of the measure theoretic framework for solving the inverse problem. In Section 2.2 we consider the numerical algorithm that simultaneously approximates the support of the probability density solving the stochastic inverse problem. We use simple linear maps to demonstrate the impact the geometry of the inverse density has on the numerical approximation of this support. In Section 3 we quantify the two guiding geometric properties of the support, the μ_Λ -*measure* and the *skewness*, which represent, in a sense, the precision and accuracy, respectively, of the

computed inverse density. In Section 4 we address the need to optimize the minimization of both the μ_Λ -measure and skewness. With this multi-criteria optimization problem defined, we use simple linear and nonlinear maps to develop an intuition about solutions to this problem. In Section 5, we demonstrate the impact of determining optimal sets of QoI on solutions to stochastic inverse problems involving physics based models. In Section 6 we provide concluding remarks and discuss promising future research directions.

2 A Measure-Theoretic Stochastic Inverse Problem

2.1 A Mathematical Formulation

We provide a taxonomy of three forward problems with increasing levels of uncertainty and the direct inverses of these problems along with a brief summary of the corresponding solutions. For a more thorough description and analysis of these problems and solutions, see [2, 5]. Let Λ denote the set of all uncertain input parameters for some physics-based model and \mathcal{D} denote the set of possible output data defined as the range of the QoI map, $Q : \Lambda \rightarrow \mathcal{D}$, which we assume is piecewise smooth.

Level 1

The first forward problem is the simplest in predictive science which is to evaluate the map Q for a fixed $\lambda \in \Lambda$ to determine the corresponding output datum $Q(\lambda) = q \in \mathcal{D}$. In other words, once the inputs of a model are specified, the simplest forward problem is to solve the model in order to predict the output datum. Generally, solutions to this problem require developing a computational model to determine approximate solutions to the physics-based model and to then apply a functional to the numerical solution to obtain the value of q . With the exception of possible discretization errors in numerically evaluating $Q(\lambda)$, there is no uncertainty in the forward problem since $Q(\lambda)$ is well-defined.

The corresponding inverse problem is to determine the possible parameters $\lambda \in \Lambda$ that produce a particular value of $q \in \mathcal{D}$. Oftentimes $Q^{-1}(q)$ defines a set of values in Λ either due to non-linearities in the map Q and/or the dimension of Λ being greater than the dimension of \mathcal{D} . Thus, the simplest inverse problem often has uncertainty as to the particular value of $\lambda \in \Lambda$ that produced a fixed output datum. However, there is no uncertainty about which set-valued inverse produced the fixed output datum. When $\dim(\Lambda) = 2$ and $\dim(\mathcal{D}) = 1$, a set-valued inverse, $Q^{-1}(q)$, defines a contour in Λ that we are familiar with from contour maps, see the left plot in Figure 1. We restrict focus to problems where $\Lambda \subset \mathbb{R}^n$, $\mathcal{D} \subset \mathbb{R}^m$, and $m < n$ (or occasionally $m \leq n$), which corresponds to the common case when there are more input parameters than observable outputs. We refer to the set-valued inverses of the map Q as *generalized contours*. Generalized contours may be defined by the union of piecewise-defined manifolds in Λ , and the local dimension of the manifolds depends upon m , n , and the rank of the Jacobian of Q .

Definition 1. *The component maps of the m -dimensional piecewise-smooth vector-*

valued map $Q(\lambda)$ are **geometrically distinct (GD)** if the Jacobian of Q has full rank at every point in Λ . When the component maps are GD, we say that Q is GD.

When $m < n$ and Q is GD, the generalized contours exist as piecewise-defined $(n - m)$ -dimensional manifolds in Λ . In [2], a method for explicit pointwise approximation of generalized contours for scalar multivariate Q maps is provided. In [5], this was extended for the more general case of vector-valued multivariate Q maps. While explicit approximations are required for numerical solutions to this “simplest” inverse problem, it is thankfully not necessary for solutions to the stochastic inverse problem which implicitly exploit the geometric structure of the generalized contour map.

Level 2

Following solutions to the first forward problem, it is typical to define a second forward problem from the general class of problems that fall under the category of deterministic sensitivity analysis. Some specific deterministic sensitivity analysis problems arise in the context of dynamical systems analysis, where we study how sets of initial conditions vary over time, or in perturbation analysis, where we study how small changes in model inputs affect model outputs. These types of forward problems can often be written mathematically as analyzing $Q(A) \subset \mathcal{D}$ given a set $A \subset \Lambda$. In other words, there is generally uncertainty as to the precise value λ takes in a set $A \subset \Lambda$ and subsequently there is uncertainty as to the particular datum to predict in the set $Q(A)$. Answering such questions often requires, at a minimum, specification of metrics on both Λ and \mathcal{D} so that we can describe distances between points. Thus, in the formulation of a forward *deterministic* sensitivity analysis problem, we often must specify metrics on both Λ and \mathcal{D} . By transfinite induction using sets in the metric topologies, we can construct the Borel σ -algebras, \mathcal{B}_Λ and $\mathcal{B}_\mathcal{D}$, on Λ and \mathcal{D} , respectively. The metrics can be used to define outer measures on the measurable spaces $(\Lambda, \mathcal{B}_\Lambda)$ and $(\mathcal{D}, \mathcal{B}_\mathcal{D})$. Then, Carathéodory’s theorem can be used to define the Hausdorff measures, which we refer to as the “volume” measures on $(\Lambda, \mathcal{B}_\Lambda)$ and $(\mathcal{D}, \mathcal{B}_\mathcal{D})$, and denote by μ_Λ and $\mu_\mathcal{D}$, respectively. Note that if the metrics are norm-induced on $\Lambda \subset \mathbb{R}^n$ and $\mathcal{D} \subset \mathbb{R}^m$, then up to a scaling constant, μ_Λ and $\mu_\mathcal{D}$ are identical to the standard Lebesgue measures on these spaces. We observe that construction of the measure spaces $(\Lambda, \mathcal{B}_\Lambda, \mu_\Lambda)$ and $(\mathcal{D}, \mathcal{B}_\mathcal{D}, \mu_\mathcal{D})$ follows directly from the specification of metrics on the sets Λ and \mathcal{D} . Therefore, we refer to the specification of metrics as the *minimal assumptions* required to obtain the measure-theoretic framework in which the remaining problems in the taxonomy are formulated.

It is worth noting that a metric need not be specified on \mathcal{D} as long as \mathcal{D} is a topological space. In this case, we can still obtain a Borel σ -algebra on \mathcal{D} as before, and define the volume measure $\mu_\mathcal{D}$ on $(\mathcal{D}, \mathcal{B}_\mathcal{D})$ in terms of the induced “push forward” measure

$$\mu_\mathcal{D}(A) = \mu_\Lambda(Q^{-1}(A)), \quad A \in \mathcal{B}_\mathcal{D}, \quad (1)$$

where the measurability of Q , which follows from the assumed smoothness properties of Q , implies $Q^{-1}(A) \in \mathcal{B}_\Lambda$.

The corresponding *deterministic* inverse sensitivity analysis problem is to analyze $Q^{-1}(A)$ for some $A \in \mathcal{B}_\mathcal{D}$. While $Q^{-1}(A) \in \mathcal{B}_\Lambda$, the practical computation of $Q^{-1}(A)$

is complicated by the fact that Q^{-1} does not map individual points in \mathcal{D} to individual points in Λ . As described above, assuming $m < n$, Q^{-1} maps a point in \mathcal{D} to an $(n - m)$ -dimensional manifold embedded in Λ . Thus, $Q^{-1}(A)$ defines a *generalized contour event*, or just contour event, belonging to an *induced contour σ -algebra* $\mathcal{C}_\Lambda \subset \mathcal{B}_\Lambda$ (where the inclusion is often proper). In other words, solutions to the corresponding inverse problem can be described in the measurable space $(\Lambda, \mathcal{C}_\Lambda)$ and the volume measure μ_Λ can be used to provide quantitative assessments of solutions to this inverse sensitivity analysis problem.

As shown below in describing solutions to the stochastic inverse problem, it is useful to consider solutions to this deterministic inverse sensitivity problem in a space of equivalence classes. Specifically, we may use the generalized contours to define an equivalence class representation of Λ where two points are considered equivalent if they lie on the same generalized contour. We let \mathcal{L} denote the space of such equivalence classes and let $\pi_\mathcal{L} : \Lambda \rightarrow \mathcal{L}$ denote the projection map where $\pi_\mathcal{L}(\lambda) = \ell \in \mathcal{L}$ defines the equivalence class corresponding to a particular λ and $\pi_\mathcal{L}^{-1}(\ell) = C_\ell$ is the generalized contour in Λ corresponding to the point $\ell \in \mathcal{L}$. It is possible to explicitly represent \mathcal{L} in Λ by choosing a specific representative element from each equivalence class. As described in [2, 5], such a representation of \mathcal{L} can be constructed by piecewise m -dimensional manifolds that *index* the $(n - m)$ -dimensional generalized contours. From a non-technical perspective, this is like defining a particular hiking path using a contour map where each elevation is transversed once and only once, see the middle plot in Figure 1. We refer to any such indexing manifold as a *transverse parameterization*. Given a particular indexing manifold representing \mathcal{L} , Q defines a bijection between \mathcal{L} and \mathcal{D} . The measure space $(\mathcal{L}, \mathcal{B}_\mathcal{L}, \mu_\mathcal{L})$ can be defined as an induced space using the bijection Q and $(\mathcal{D}, \mathcal{B}_\mathcal{D}, \mu_\mathcal{D})$. Then, solutions to the deterministic inverse sensitivity analysis problem can be described and analyzed as measurable sets of points in \mathcal{L} instead of measurable generalized contour events in Λ . While the measure space $(\mathcal{L}, \mathcal{B}_\mathcal{L}, \mu_\mathcal{L})$ is useful in describing both the theoretical solutions and computational algorithms approximating solutions to the stochastic inverse problem defined below, it is not necessary to explicitly construct a transverse parameterization.

Level 3

The third forward problem builds on the second forward problem. We now assume that there is uncertainty, described in terms of probabilities, as to which set $A \in \mathcal{B}_\Lambda$ the model parameters λ belong, and the goal is to analyze the probabilities of sets in $\mathcal{B}_\mathcal{D}$. In the language of probability theory, measurable sets are referred to as events. Thus, we assume that a probability measure P_Λ is given on $(\Lambda, \mathcal{B}_\Lambda)$ describing uncertainty in the events for which parameters may belong, and the goal is to determine $P_\mathcal{D}$ on $(\mathcal{D}, \mathcal{B}_\mathcal{D})$. We refer to this as the *stochastic forward problem*. When P_Λ (resp., $P_\mathcal{D}$) is absolutely continuous with respect to the volume measure μ_Λ (resp., $\mu_\mathcal{D}$), the corresponding Radon-Nikodym derivative (i.e., the probability density function) ρ_Λ (resp., $\rho_\mathcal{D}$) is usually given in place of the probability measure. We assume this is the case so that we can refer to probabilities of events in terms of the more common representation using integrals and

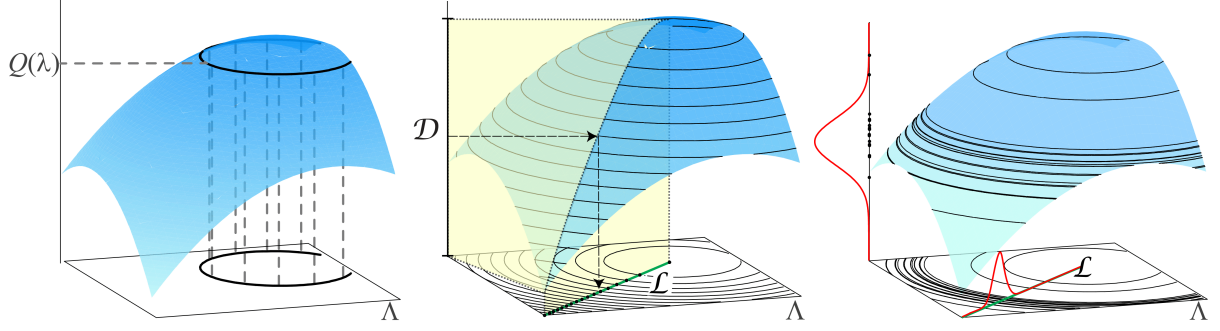


Figure 1: *Left: The inverse of $Q^{-1}(Q(\lambda))$ is often set-valued even when $\lambda \in \Lambda$ specifies a particular $Q(\lambda) \in \mathcal{D}$. Middle: The representation of \mathcal{L} as a transverse parameterization. Right: A probability measure described as a density on $(\mathcal{D}, \mathcal{B}_{\mathcal{D}})$ maps uniquely to a probability density on $(\mathcal{L}, \mathcal{B}_{\mathcal{L}})$. Figures adopted from [5] and [2].*

density functions, e.g.,

$$P_{\Lambda}(A) = \int_A \rho_{\Lambda} d\mu_{\Lambda}, \quad A \in \mathcal{B}_{\Lambda}. \quad (2)$$

Solution to this stochastic forward problem are given by the induced push-forward probability measure $P_{\mathcal{D}}$ defined for any $A \in \mathcal{B}_{\mathcal{D}}$ by

$$P_{\mathcal{D}}(A) = \int_A \rho_{\mathcal{D}} d\mu_{\mathcal{D}} = \int_{Q^{-1}(A)} \rho_{\Lambda} d\mu_{\Lambda} = P_{\Lambda}(Q^{-1}(A)). \quad (3)$$

This is a familiar problem in uncertainty quantification. The approximate solution can be obtained by a classic Monte Carlo method.

The corresponding inverse problem assumes we are given a probability measure $P_{\mathcal{D}}$ on $(\mathcal{D}, \mathcal{B}_{\mathcal{D}})$, and the goal is to determine a probability measure P_{Λ} on $(\Lambda, \mathcal{B}_{\Lambda})$ such that $P_{\Lambda}(Q^{-1}(A)) = P_{\mathcal{D}}(A)$ for all $A \in \mathcal{B}_{\mathcal{D}}$. In other words, the push-forward measure of a solution to the stochastic inverse problem should be $P_{\mathcal{D}}$. To determine such a solution to the stochastic inverse problem, we first consider the stochastic inverse problem posed on $(\mathcal{L}, \mathcal{B}_{\mathcal{L}})$.

Since Q defines a bijection between \mathcal{L} and \mathcal{D} , there is a unique $P_{\mathcal{L}}$ on $(\mathcal{L}, \mathcal{B}_{\mathcal{L}})$ (see the right-hand plot in Figure 1 and [5]). We can then use the projection map $\pi_{\mathcal{L}}$ to prove the following theorem.

Theorem 1. *The stochastic inverse problem has a unique solution on $(\Lambda, \mathcal{C}_{\Lambda})$.*

However, the goal is to define a probability measure P_{Λ} on the measurable space $(\Lambda, \mathcal{B}_{\Lambda})$ not on a space involving contour events that are complicated to describe. This requires an application of the Disintegration Theorem, which allows for the rigorous description of conditional probability measures defined on sets of zero μ_{Λ} -measure [5, 6, 9].

Theorem 2. *Let $(\Lambda, \mathcal{B}_{\Lambda})$ be a measurable space. Assume that P_{Λ} is a probability measure on $(\Lambda, \mathcal{B}_{\Lambda})$. There exists a family of conditional probability measures $\{P_{\ell}\}$ on $\{(C_{\ell}, \mathcal{B}_{C_{\ell}})\}$*

giving the disintegration,

$$P_\Lambda(A) = \int_{\pi_{\mathcal{L}}(A)} \left(\int_{\pi_{\mathcal{L}}^{-1}(\ell) \cap A} dP_\ell(\lambda) \right) dP_{\mathcal{L}}(\ell), \quad \forall A \in \mathcal{B}_\Lambda. \quad (4)$$

Thus, any probability measure on $(\Lambda, \mathcal{B}_\Lambda)$ can be decomposed into a form involving a probability measure on $(\mathcal{L}, \mathcal{B}_{\mathcal{L}})$ uniquely defined by $P_{\mathcal{D}}$ and probability measures on each measurable generalized contour space $(C_\ell, \mathcal{B}_{C_\ell})$ defined by the conditional probabilities P_ℓ . The conditional probability measures on $\{(C_\ell, \mathcal{B}_{C_\ell})\}$ can not be determined by observations of $Q(\lambda) \in \mathcal{D}$. We follow [5, 6] and adopt what is referred to as the *standard Ansatz* determined by the disintegration of the volume measure μ_Λ to compute probabilities of events inside of a contour event. The standard Ansatz is given by

$$P_\ell = \mu_{C_\ell} / \mu_{C_\ell}(C_\ell), \quad \forall \ell \in \mathcal{L}, \quad (5)$$

where μ_{C_ℓ} is the disintegrated volume measure on the generalized contour C_ℓ .

Note that the standard Ansatz can be used as long as $\mu_\Lambda(\Lambda) < \infty$. The approximation method and resulting non-intrusive computational algorithm can be easily modified for any Ansatz. See [5] for more details and theory regarding general choices of the Ansatz. Combining an Ansatz with the Disintegration Theorem proves the following theorem.

Theorem 3. *Under the Ansatz, the stochastic inverse problem has a unique solution on $(\Lambda, \mathcal{B}_\Lambda)$.*

The standard Ansatz results in a probability measure P_Λ that inherits key geometric features from the generalized contour map. The only assumption necessary for employing the standard Ansatz is that a finite volume measure exists, which actually means we only assume that a metric exists as described above and that the diameter of Λ is bounded. Any other choice of Ansatz imposes some other geometric constraints on the solution to the inverse problem that are not coming from the map Q . If we choose not to use any Ansatz, then we can always solve the stochastic inverse problem on $(\Lambda, \mathcal{C}_\Lambda)$, i.e., we can always compute the unique probability measure when restricted to contour events.

2.2 Numerical Approximation of Solutions

Fundamental to approximating solutions P_Λ to the stochastic inverse problem is the approximation of events in the various σ -algebras, $\mathcal{B}_{\mathcal{D}}$, \mathcal{C}_Λ and \mathcal{B}_Λ . Since $\mathcal{C}_\Lambda \subset \mathcal{B}_\Lambda$, we can simultaneously approximate events in both of these σ -algebras using the same set of events partitioning Λ . Let $\{\mathcal{V}_i\}_{i=1}^N$ denote such a partition of Λ where $\mathcal{V}_i \in \mathcal{B}_\Lambda$ for each i . Assume that we are given a collection of sets $\{D_k\}_{k=1}^M \subset \mathcal{B}_{\mathcal{D}}$ partitioning \mathcal{D} . The basic algorithmic procedure for approximating $P_\Lambda(\mathcal{V}_i)$ for each i is to determine which of the $\{\mathcal{V}_i\}_{i=1}^N$ s approximate $Q^{-1}(D_k)$, and then to apply the Ansatz on this approximation of the contour event which has known probability $P_{\mathcal{D}}(D_k)$. Letting $p_{\Lambda,i}$ denote this approximation, we can then define an approximation to P_Λ as

$$P_\Lambda(A) \approx P_{\Lambda,N}(A) = \sum_{i=1}^N p_{\Lambda,i} \chi_{\mathcal{V}_i}(A), \quad A \in \mathcal{B}_\Lambda.$$

We summarize this basic procedure for approximating P_Λ in Algorithm 1.

Algorithm 1: Numerical Approximation of Inverse Probability Measure

1. Let $\{\mathcal{V}_i\}_{i=1}^N \subset \mathcal{B}_\Lambda$ partition Λ .
 2. Determine a nominal value Q_i for the map $Q(\lambda)$ on \mathcal{V}_i for $i = 1, \dots, N$.
 3. Choose a partitioning of \mathcal{D} , $\{D_k\}_{k=1}^M \subset \mathcal{D}$.
 4. Compute $p_{\mathcal{D},k} = P_{\mathcal{D}}(D_k)$ for $k = 1, \dots, M$.
 5. Let $\mathcal{C}_k = \{i | Q_i \in D_k\}$ for $k = 1, \dots, M$.
 6. Let $\mathcal{O}_i = \{k | Q_i \in D_k\}$ for $i = 1, \dots, N$.
 7. Compute $V_i = \mu_\Lambda(\mathcal{V}_i)$ for $i = 1, \dots, N$.
 8. Set $p_{\Lambda,i} = (V_i / \sum_{j \in \mathcal{C}_{\mathcal{O}_i}} V_j) p_{\mathcal{D},\mathcal{O}_i}$ for $i = 1, \dots, N$.
-

In Algorithm 1, \mathcal{C}_k is used to determine which sets from $\{\mathcal{V}_i\}_{i=1}^N$ approximate the contour event $Q^{-1}(D_k)$ for each k . Similarly, \mathcal{O}_i is used to determine which contour event $Q^{-1}(D_k)$ is associated to \mathcal{V}_i for each i . The Ansatz is applied in the final step where the probability of each \mathcal{V}_i is determined by the probability $P_{\mathcal{D}}(D_{\mathcal{O}_i})$ multiplied by the ratio of the volume of \mathcal{V}_i to the volume of the approximate contour event of $Q^{-1}(D_{\mathcal{O}_i})$.

The ability to accurately approximate contour events by subsets of $\{\mathcal{V}_i\}_{i=1}^N$ is critical in computing accurate approximations of P_Λ . In other words, the ability to numerically approximate solutions to the deterministic inverse sensitivity problem determines the numerical accuracy in the stochastic inverse problem. Refining the partition $\{\mathcal{V}_i\}_{i=1}^N$ generally means additional solutions of the model are required in order to compute the nominal value Q_i required in Step 2 of Algorithm 1. A typical goal is to determine the minimal number of sets partitioning Λ such that P_Λ can be approximated for a fixed discretization of $P_{\mathcal{D}}$. Such ideas and goals were first explored in [7] where the concept of *skewness* was defined, which described the computational complexity in solving the inverse problem entirely in terms of the number of sets partitioning Λ required for accurate approximation of P_Λ . While there are other approximation issues that arise from numerical evaluation of the map Q and approximation of measures by densities on partitions of the various spaces, these have been addressed elsewhere, e.g., see [4, 6], and are only exacerbated by the fundamental problem of approximating contour events. To illustrate the geometric concept of skewness and the affect on the number of samples N needed in Algorithm 1 to accurately approximate the inverse image, we use the simple example below, where for simplicity, we assume $n = m$.

Suppose we are given two different QoI maps $Q^{(a)} = (Q_1^{(a)}, Q_2^{(a)}) : \Lambda = [0, 1]^2 \rightarrow \mathcal{D}^{(a)} \subset$

\mathbb{R}^2 and $Q^{(b)} = (Q_1^{(b)}, Q_2^{(b)}) : \Lambda = [0, 1]^2 \rightarrow \mathcal{D}^{(b)} \subset \mathbb{R}^2$ defined by

$$Q_1^{(a)}(\lambda_1, \lambda_2) = \lambda_1, \quad (6)$$

$$Q_2^{(a)}(\lambda_1, \lambda_2) = \lambda_2, \quad (7)$$

$$Q_1^{(b)}(\lambda_1, \lambda_2) = \lambda_1 + \lambda_2, \quad (8)$$

$$Q_2^{(b)}(\lambda_1, \lambda_2) = 0.74\lambda_1 + 1.26\lambda_2. \quad (9)$$

Let $\mathcal{D}^{(a)}$ and $\mathcal{D}^{(b)}$ be partitioned by rectangles $\left\{D_k^{(a)}\right\}_{k=1}^{M_a}$ and $\left\{D_l^{(b)}\right\}_{l=1}^{M_b}$, respectively. In Figure 2, we show the affect of skewness by considering $(Q^{(a)})^{-1}(D_k^{(a)})$ and $(Q^{(b)})^{-1}(D_l^{(b)})$ for a fixed k and l , the error in the Voronoi approximation of this contour event which is given by the symmetric difference (shown by the shaded regions of green), and how this error decreases with increased N . In Figure 3 and Table 1, we show the convergence of the measure of the symmetric difference to be near the known convergence rate of MC methods, $N^{-1/2}$. Although numerical approximations of the inverse densities converge at the same rate for both $Q^{(a)}$ and $Q^{(b)}$, for a given N the error for the well conditioned map $(Q^{(a)})$ is about half the error of the poorly conditioned (or highly skewed) map $(Q^{(b)})$. We describe how to quantify the skewness of a given map Q in Section 3.2.

3 Precision and Accuracy

Following solution to the stochastic inverse problem, we generally want to make, and quantify uncertainty in, predictions that can be formulated as solutions to a stochastic forward problem, e.g., computing a confidence interval on storm surge levels of a hurricane forecast using an ensemble of predictions determined from sampling model input parameters. In some cases, we may be able to design the observation network defining the QoI map used in the stochastic inverse problem, e.g., by determining where and when to deploy sensors such as buoys in the Gulf of Mexico. The problem of defining the “best” observation network is generally referred to as optimal experimental design. Below, we motivate both the definitions we use to define what is an optimal QoI map as well as the development of criteria for choosing the optimal QoI map. Our starting point is the assumption that we solve the stochastic inverse problem in order to *quantifiably improve* the predictive capabilities of the model.

Suppose two different QoI maps, $Q^{(a)}$ and $Q^{(b)}$, define two stochastic inverse problems with solutions represented by the two densities $\rho_\Lambda^{(a)}$ and $\rho_\Lambda^{(b)}$ such that

$$\mu_\Lambda(\text{supp}(\rho_\Lambda^{(a)})) \ll \mu_\Lambda(\text{supp}(\rho_\Lambda^{(b)})).$$

For simplicity, assume that $\Lambda \subset \mathbb{R}^n$, μ_Λ is the Lebesgue measure, and the supports of the densities are convex subsets. Let $Q^{(p)}$ denote a prediction QoI that is a linear map on Λ , and let $\mathcal{D} = Q^{(p)}(\Lambda)$ denote the space of possible outcomes for the predictions. Then, $Q^{(p)}(\text{supp}(\rho_\Lambda^{(a)}))$ and $Q^{(p)}(\text{supp}(\rho_\Lambda^{(b)}))$ define the events of all probable predictions based

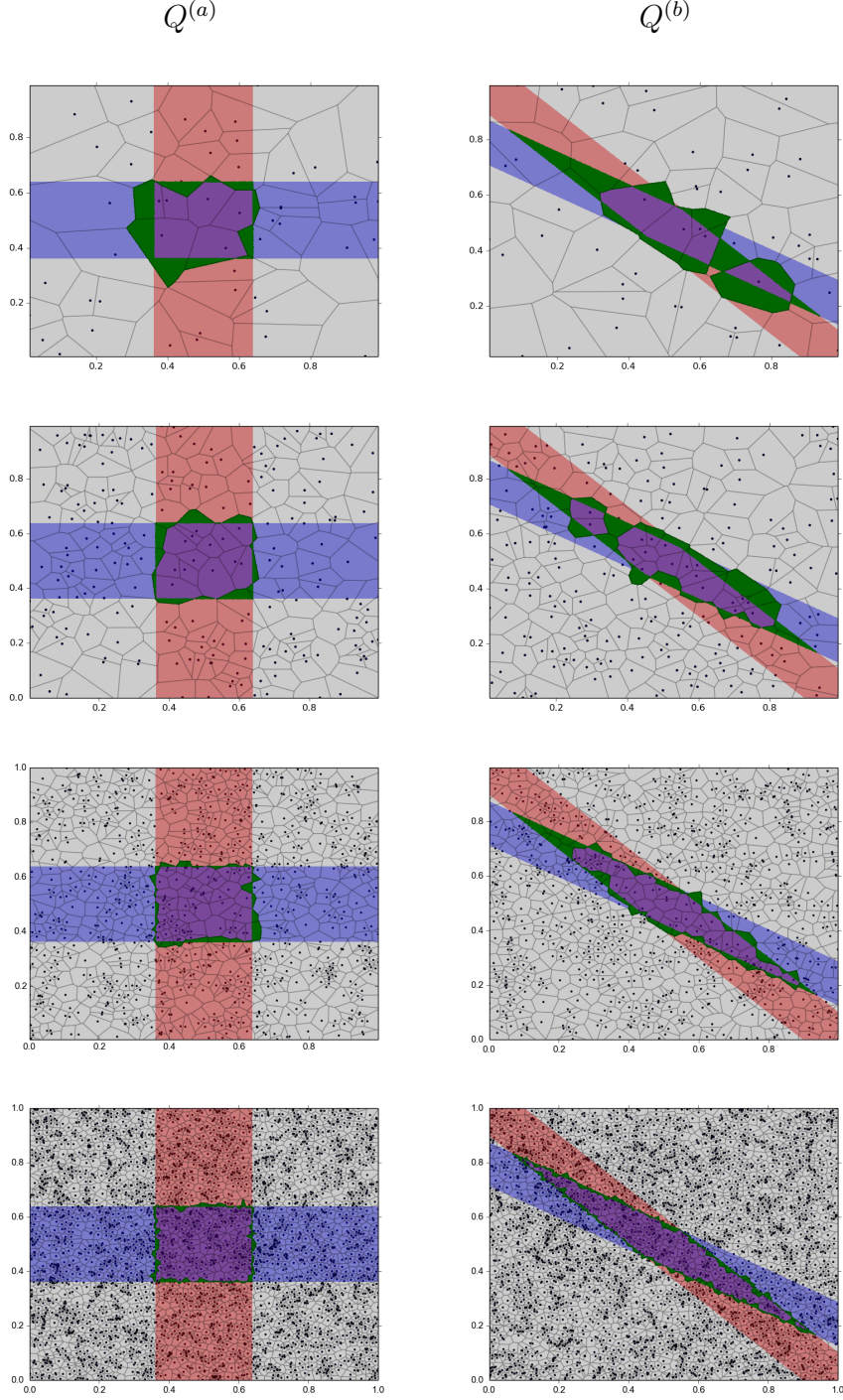


Figure 2: We plot the implicit discretization of $\Lambda \{\mathcal{V}_i\}_{i=1}^N$ (given by the uniform random samples), the exact inverse density (given by the intersection of the red and blue contour events) of $D_k^{(a)}$ and $D_l^{(b)}$ for fixed k and l , and the symmetric difference of the exact inverse image with the approximated inverse image (seen in green). We show this for both $Q^{(a)}$ (left) and $Q^{(b)}$ (right). This is shown for 50 samples (top), 200 samples (second from the top), 800 samples (second from the bottom), and 3200 samples (bottom).

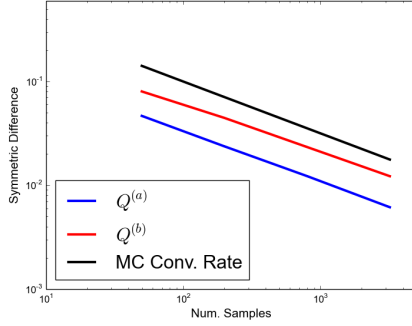


Figure 3: *Loglog convergence plot for the mean symmetric difference shown in Table 1.*

Num. Samples	$Q^{(a)}$	$Q^{(b)}$
50	$4.66E-2$	$8.01E-2$
200	$2.37E-2$	$4.45E-2$
800	$1.22E-2$	$2.33E-2$
3200	$6.13E-3$	$1.22E-2$

Table 1: *Mean of the measure of the symmetric difference for 100 sets of N random samples.*

on solutions to the separate stochastic inverse problems, and $\mu_{\mathcal{D}}(Q^{(p)}(\text{supp}(\rho_{\Lambda}^{(a)}))) \ll \mu_{\mathcal{D}}(Q^{(p)}(\text{supp}(\rho_{\Lambda}^{(b)})))$. Consequently, we generally expect that statistical inferences drawn from $Q^{(p)}(\text{supp}(\rho_{\Lambda}^{(a)}))$ (e.g., the mean of the predictions) to have greater precision in terms of reduced variances, smaller confidence intervals, etc., compared to those drawn from $Q^{(p)}(\text{supp}(\rho_{\Lambda}^{(b)}))$.

This motivates a general measure-theoretic goal for designing experiments where the goal is to determine an observation network defining the map Q such that solution to the stochastic inverse problem results in a significant amount of probability contained in events of small μ_{Λ} -measure. It is clear from Eq. (4), the Ansatz, and Algorithm 1, that the μ_{Λ} -measures of induced contour events play a pivotal role in determining an optimal map Q . In Section 3.1, we describe an efficient computational approach for quantifying the μ_{Λ} -measures of induced contour events for nonlinear maps Q . Given multiple choices for the QoI map, we may use such results to determine the so-called optimal map Q . However, given finite computational resources, it may not be possible to accurately approximate solutions of the stochastic inverse problem using the optimal map Q if only a relatively small number of model solves are allowed as described in Section 2.2. A computational method for quantifying the global affect of skewness from a nonlinear map Q is described in Section 3.2. In Section 4, we describe how to optimize the choice of Q to take into account the separate goals of obtaining *precise* predictions from stochastic inverse problems that can be solved *accurately* with relatively few model evaluations.

3.1 μ_{Λ} -measure of support

Suppose in an experiment we can obtain a total of m QoI. Let $B \in \mathcal{B}_{\mathcal{D}}$ define a typical output event from a partition of \mathcal{D} used in Algorithm 1. If $P_{\mathcal{D}}(B) \approx 1$, then as described above, the goal is to quantify the expected μ_{Λ} -measure of the support of $Q^{-1}(B)$. Since we do not know $P_{\mathcal{D}}$ a priori, we want to choose a B that is representative of typical

tessellations used to discretize any probability measure and also geometrically easy to describe such that $\mu_\Lambda(Q^{-1}(B))$ is computationally inexpensive to approximate. In measure theory, it is common for σ -algebras on higher dimensional spaces to be generated from generalized rectangles. We therefore consider B to be a generalized rectangle in $\mathcal{D} \subset \mathbb{R}^m$. Suppose Q is a linear GD map and $n = m$, i.e., Q is defined by the square invertible matrix J , $Q(\lambda) = J\lambda$. In this case, if $\Lambda = \mathbb{R}^n$, we have that

$$\mu_\Lambda(Q^{-1}(B)) = \mu_\Lambda(J^{-1}(B)) = \mu_{\mathcal{D}}(B) \det(J^{-1}) = \frac{\mu_{\mathcal{D}}(B)}{\det J}. \quad (10)$$

Note that if $\Lambda \subset \mathbb{R}^n$ is proper, then the above equation is not necessarily true as $Q^{-1}(B)$ may intersect the boundary of Λ . We neglect such boundary effects in the computations, and simply note that in certain cases they may play an important role although this is not the typical case in our experience.

In general Q is not linear, so we consider local linear approximations of Q to quantify the μ_Λ -measure of inverting sets such as B into different regions of Λ . Given $\lambda^{(i)} = (\lambda_1^{(i)}, \lambda_2^{(i)}, \dots, \lambda_n^{(i)}) \in \Lambda$, denote the Jacobian of Q evaluated at $\lambda^{(i)}$, by

$$J_{\lambda^{(i)}} = \begin{bmatrix} \frac{\partial Q_1(\lambda^{(i)})}{\partial \lambda_1} & \cdots & \frac{\partial Q_1(\lambda^{(i)})}{\partial \lambda_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial Q_m(\lambda^{(i)})}{\partial \lambda_1} & \cdots & \frac{\partial Q_m(\lambda^{(i)})}{\partial \lambda_n} \end{bmatrix}. \quad (11)$$

For simplicity in describing the sets $Q^{-1}(B)$ we initially assume $n = m$ and the map Q is GD, which implies that the Jacobian is square and invertible. Suppose now that the generalized rectangle B is centered at $Q(\lambda^{(i)})$, then from above we have that

$$\mu_\Lambda(Q^{-1}(B)) \approx M_Q(\lambda^{(i)}) := \mu_\Lambda(J_{\lambda^{(i)}}^{-1}(B)) = \frac{\mu_{\mathcal{D}}(B)}{\det J_{\lambda^{(i)}}}. \quad (12)$$

Here, we introduce the notation $M_Q(\lambda^{(i)})$ as a shorthand for the size (i.e., measure) of local induced contour events defined around a sample $\lambda^{(i)}$, and we make explicit the dependence on the QoI map Q . Since the determinant of J can be written as the product of the singular values of J , we have

$$M_Q(\lambda^{(i)}) = \mu_{\mathcal{D}}(B) \prod_{k=1}^m \sigma_{ik}^* = \frac{\mu_{\mathcal{D}}(B)}{\prod_{k=1}^m \sigma_{ik}}, \quad (13)$$

where $\{\sigma_{ik}^*\}_{k=1}^m$ are the singular values of $J_{\lambda^{(i)}}^{-1}$ and $\{\sigma_{ik}\}_{k=1}^m$ are the singular values of $J_{\lambda^{(i)}}$. The *average* $M_Q(\lambda)$ is given by

$$\overline{M_Q} = \frac{1}{\mu_\Lambda(\Lambda)} \int_\Lambda M_Q(\lambda) d\mu_\Lambda. \quad (14)$$

Given a set of N samples $\{\lambda^{(i)}\}_{i=1}^N$ in Λ , we approximate $\overline{M_Q}$ using the Monte Carlo estimate

$$\overline{M_Q} \approx \overline{M_{Q,N}} := \frac{1}{N} \sum_{i=1}^N M_Q(\lambda^{(i)}) = \frac{1}{N} \sum_{i=1}^N \frac{\mu_{\mathcal{D}}(B)}{\prod_{k=1}^m \sigma_{ik}}. \quad (15)$$

We now relax the assumption that $n = m$ and consider the (more common) case of $m < n$. In this case, we recall that a transverse parameterization, which is an explicit representation in Λ of the equivalence class structure \mathcal{L} defined by the generalized contours, exists as a (possibly piecewise defined) m -dimensional manifold. Restricting $Q : \Lambda \rightarrow \mathcal{D}$ to be $Q : \mathcal{L} \rightarrow \mathcal{D}$, which we denote by $Q|_{\mathcal{L}}$, we then have an $m \times m$ map and can apply the above computations to $Q|_{\mathcal{L}}$. In other words, we are now interested in the $\mu_{\mathcal{L}}$ -measure of $Q|_{\mathcal{L}}^{-1}(B) \in \mathcal{B}_{\mathcal{L}}$. Geometrically, we may think of this problem as determining the measure of the *cross section of the contour event intersected with a transverse parameterization manifold* in Λ . Given a local linear approximation of Q at some point $\lambda^{(i)} \in \Lambda$, the following lemma shows that we may approximate $\mu_{\mathcal{L}}(M_{Q|_{\mathcal{L}}}(\lambda^{(i)}))$ similar to before. Consequently, the lemma implies that we may compute $\overline{M}_{Q|_{\mathcal{L}}, N}$ using $M_{Q|_{\mathcal{L}}}(\lambda^{(i)})$ in Eq. (15).

Lemma 1. *Let J be a full rank $m \times n$ matrix with $m \leq n$. Then the Lebesgue measure μ in \mathbb{R}^m of the m -dimensional parallelepiped defined by the m rows of J , denoted $\mu(Pa(J))$, is given by the product of the singular values of J ,*

$$\mu(Pa(J)) = \prod_{i=1}^m \sigma_i. \quad (16)$$

Proof. The singular values of J are equal to the singular values of J^\top . Consider the reduced QR factorization of J^\top ,

$$J^\top = \tilde{Q}R, \quad (17)$$

where \tilde{Q} is $n \times m$ and R is $m \times m$. By the properties of the QR factorization, we know the singular values of R are the same as the singular values of J^\top . Let $x \in \mathbb{R}^m$, then

$$\|\tilde{Q}x\|^2 = (\tilde{Q}x)^\top (\tilde{Q}x) = x^\top \tilde{Q}^\top \tilde{Q}x = x^\top x = \|x\|^2, \quad (18)$$

so \tilde{Q} is an isometry. This implies the Lebesgue measure of the parallelepiped defined by the rows of R is equal to the Lebesgue measure of the parallelepiped defined by the columns of J^\top , or the rows of J ,

$$\mu(Pa(J)) = \mu(Pa(R)) = \prod_{k=1}^m \gamma_k = \prod_{k=1}^m \sigma_k, \quad (19)$$

where γ_k are the singular values of R and σ_k are the singular values of J . \square

3.2 Skewness

While determining the QoI map Q that optimally reduces $\mu_{\Lambda}(Q^{-1}(B))$ for $B \in \mathcal{B}_{\mathcal{D}}$ should reduce uncertainty (i.e., increase precision) in predictions, we often need to computationally approximate $Q^{-1}(B)$ as described in Algorithm 1. Errors in the approximations to P_{Λ} propagate to errors in predictions, and while the predicted sets of high probability may have small $\mu_{\mathcal{D}}$ -measure, the errors in the prediction may be so large as to render

such predictions useless in practice. Hence, in this section we quantify the *skewness* of local induced contour maps and then expand this to the *average skewness* as done in Section 3.1 for the μ_Λ -measure, M_Q .

Again, we refer to local linear approximations of the given map Q at some $\lambda^{(i)} \in \Lambda$, $J_{\lambda^{(i)}}$. Let $Pa(J_{\lambda^{(i)}})$ denote the parallelepiped defined by the n -dimensional vectors j_1, \dots, j_m that are the rows of $J_{\lambda^{(i)}}$. We use $\mu(Pa(J_{\lambda^{(i)}}))$ to denote the measure of the parallelepiped as an m -dimensional object. This measure is calculated in terms of the singular values of $J_{\lambda^{(i)}}$ by Lemma 1. A fundamental decomposition is

Theorem 4. *Given n -dimensional vectors j_1, \dots, j_m , there exists vectors j_1^\perp, j_1^0 such that*

$$j_1 = j_1^\perp + j_1^0, \quad j_1^\perp \perp j_1^0, \quad j_1^0 \in \text{span}\{j_2, \dots, j_m\},$$

and

$$\mu(Pa(J_{\lambda^{(i)}})) = |j_1^\perp| \times \mu(Pa(J_{1,\lambda^{(i)}})), \quad (20)$$

where $J_{\lambda^{(i)}}$ denotes the $m \times n$ matrix formed by the vectors j_1, \dots, j_m , $J_{1,\lambda^{(i)}}$ denotes the matrix $J_{\lambda^{(i)}}$ with the first row deleted, and $\mu(Pa(J_{1,\lambda^{(i)}}))$ denotes the measure of the parallelepiped defined by the n -dimensional vectors j_2, \dots, j_m as an $(m-1)$ -dimensional object.

With this fundamental decomposition we define the local skewness of a given map Q at some point $\lambda^{(i)} \in \Lambda$, see [7].

Definition 2. *For a vector j_k , we define*

$$S_Q(J_{\lambda^{(i)}}, j_k) = \frac{|j_k|}{|j_k^\perp|}. \quad (21)$$

Then we define the **local skewness** at a point $\lambda^{(i)}$ as

$$S_Q(J_{\lambda^{(i)}}) = \max_k S_Q(J_{\lambda^{(i)}}, j_k). \quad (22)$$

Corollary 1. *The local skewness of Q , $S_Q(J_{\lambda^{(i)}})$, can be completely determined by the the norms of n -dimensional vectors and products of singular values of the Jacobians of QoI maps of dimensions $m-1$ and m ,*

$$S_Q(J_{\lambda^{(i)}}) = \max_k \frac{|j_k| \mu(Pa(J_{k,\lambda^{(i)}}))}{\mu(Pa(J_{\lambda^{(i)}}))}. \quad (23)$$

Proof.

$$S_Q(J_{\lambda^{(i)}}) = \max_k S_Q(J_{\lambda^{(i)}}, j_k) = \max_k \frac{|j_k|}{|j_k^\perp|} = \max_k \frac{|j_k| \mu(Pa(J_{k,\lambda^{(i)}}))}{\mu(Pa(J_{\lambda^{(i)}}))}, \quad (24)$$

then applying Lemma 1 we have

$$\max_k \frac{|j_k| \mu(Pa(J_{k,\lambda^{(i)}}))}{\mu(Pa(J_{\lambda^{(i)}}))} = \max_k \frac{|j_k| \prod_{r=1}^{m-1} \sigma_{kr}}{\prod_{r=1}^m \sigma_r}. \quad (25)$$

where σ_r are the singular values of $J_{\lambda^{(i)}}$ and σ_{kr} are the singular values of $J_{k,\lambda^{(i)}}$. \square

Corollary 1 allows us to take advantage of efficient singular value decompositions to algorithmically approximate $S_Q(J_{\lambda^{(i)}})$. The average skewness is given by

$$\overline{S_Q} = \frac{1}{\mu_\Lambda(\Lambda)} \int_\Lambda S_Q(\lambda) d\mu_\Lambda, \quad (26)$$

and approximated by

$$\overline{S_Q} \approx \overline{S_{Q,N}} := \frac{1}{N} \sum_{i=1}^N S_Q(\lambda^{(i)}). \quad (27)$$

Remark 1. Notice the definition of average local skewness is independent of $B \in \mathcal{B}_\mathcal{D}$. This is because B is assumed to be a generalized rectangle, i.e., has perfect skewness properties itself. This is a reasonable assumption as B is typically defined by the cross product of intervals defining the uncertainty in each QoI.

4 Multicriteria Optimization

A common problem in designing observation networks is to determine where and when to deploy a finite number of sensors in order to record *useful* data related to a physics-based model. The various possible configurations of sensors defines a family of possible QoI maps. Given a family of possible QoI maps, our goal is to determine a particular QoI map resulting in *precise* and *accurate* numerical approximation of the inverse solution. Here, we frame the problem of determining such a QoI map as an optimization problem that simultaneously reduces both $\overline{M_Q}$ and $\overline{S_Q}$.

Let \mathcal{Q} denote the space of all possible sets of QoI, let $Q^{(z)} \in \mathcal{Q}$ represent a set of m QoI, and let $\mathcal{D}^{(z)}$ represent the m -dimensional data space defined by $Q^{(z)}$, i.e., $Q^{(z)} : \Lambda \subset \mathbb{R}^n \rightarrow \mathcal{D}^{(z)} \subset \mathbb{R}^m$. For each $Q^{(z)}$ there is a corresponding generalized rectangle $B^{(z)} \in \mathcal{B}_{\mathcal{D}^{(z)}}$ that represents the uncertainty in a possible recorded datum in $\mathcal{D}^{(z)}$.

Recall from Section 3 the goals are to reduce $\overline{S_{Q^{(z)}}}$ and $\overline{M_{Q^{(z)}}}$. Let $\mathcal{S} \subset \mathbb{R}$ denote the set of all possible values of $\overline{S_{Q^{(z)}}}$ and $\mathcal{M} \subset \mathbb{R}$ denote the set of all possible values of $\overline{M_{Q^{(z)}}}$. We then define the metric spaces $(\mathcal{S}, d_\mathcal{S})$ and $(\mathcal{M}, d_\mathcal{M})$ where

$$d_\mathcal{S}(x, y) = \frac{|x - y|}{1 + |x - y|} \text{ for all } x, y \in \mathcal{S}, \quad (28)$$

and

$$d_\mathcal{M}(x, y) = \frac{|x - y|}{1 + |x - y|} \text{ for all } x, y \in \mathcal{M}. \quad (29)$$

We define the metrics this way so that the effect of the scaling of each component is limited in solving the minimization problem defined below. Let Y_ω denote the 2-dimensional product space $\mathcal{S} \times \mathcal{M}$, with metric defined by

$$d_{Y_\omega}(x, y) = \omega d_\mathcal{S}(x_1, y_1) + (1 - \omega) d_\mathcal{M}(x_2, y_2) \text{ for all } x, y \in Y_\omega, \quad (30)$$

where $\omega \in [0, 1]$ determines the relative importance we place on either precision or accuracy.

Remark 2. *The weighting of the μ_Λ -measure and the skewness is a current topic of discussion. When more computational resources are available for solving the model, it is likely we weight the skewness less than the μ_Λ -measure. When we can only afford few samples we weight the skewness more than μ_Λ -measure. This is a topic for future work.*

With this notion of distance on Y_ω we define the multicriteria optimization problem as finding the solution to,

$$\min_{Q^{(z)} \in \mathcal{Q}} d_{Y_\omega}(p, y_z), \quad (31)$$

where $p = (1, 0)$ is the *ideal point* and $y_z = (\overline{S_{Q^{(z)}}}, \overline{M_{Q^{(z)}}})$.

4.1 Discrete Optimization

Suppose \mathcal{Q} is a finite family of possible QoI maps, e.g., as defined by identifying a finite set of d physically possible configurations of m sensors with $d \geq m$. The solution to the minimization problem defined by Eq. (31) can be found by an exhaustive search through the $\binom{d}{m}$ possible QoI maps. This combinatorial problem can clearly become computationally expensive as the number of possible maps gets large. However, it is completely straightforward to implement and is embarrassingly parallel. In the concluding remarks of Section 6, we describe some possible future directions to mitigate the cost of solving the optimization problem. We provide two numerical examples to clarify this method of determining an optimal QoI map to use in the inverse problem.

Remark 3. *We note that another equivalent way to frame the discrete optimization problem is to define a “theoretical” QoI map $Q : \Lambda \rightarrow \mathcal{D} \subset \mathbb{R}^d$ where the goal is to determine the “practical” QoI map given by a subset of m components of the map Q . The advantage of this approach from an implementation point of view is that for each sample in Λ , the model is solved once in order to compute all the possible QoI values. It also provides an index to each possible QoI making the description of the optimal QoI more straightforward. Therefore, we use this in the numerical examples below.*

4.2 Numerical Examples

4.2.1 Linear Example

Let $Q : \mathbb{R}^2 \rightarrow \mathbb{R}^3$,

$$Q = \begin{bmatrix} 0.5 & 0.5 \\ 2.5 & 0.5 \\ -0.2 & 0.3 \end{bmatrix}. \quad (32)$$

Notice the rows of Q are pairwise linearly independent, i.e., pairwise GD. Let $B^{(z)} \in \mathcal{B}_{\mathcal{D}^{(z)}}$ be the rectangle of uncertainty in the data space defined by the z^{th} possible pair of QoI. For simplicity, we let the uncertainty in each QoI be the same, i.e., $\mu_{\mathcal{D}^{(z)}}(B^{(z)})$ is constant for all z . Note that because Q is linear we can easily use the exact Jacobian of Q .

Furthermore, we use the exact averages $\overline{M_{Q(z)}}$ and $\overline{S_{Q(z)}}$ not the approximate averages. The linearity of Q implies $M_{Q(z)}(\lambda^{(i)})$ and $S_{Q(z)}(\lambda^{(i)})$ are each constant for all i , so we need not numerically approximate the integrals in Eqs. (14) and (26). In Section 4.2.2 we consider a nonlinear map and compute approximations of the Jacobian, $\overline{M_{Q(z)}}$, and $\overline{S_{Q(z)}}$.

We see in Figure 4 and Table 2 the optimal choice of pair of QoI to use in the inverse problem for three different optimization problems; minimize $\overline{M_{Q(z)}}$, minimize $\overline{S_{Q(z)}}$, and minimize Eq. (30) with $\omega = 0.5$. In the top row of Figure 4 we see the inverse image of $B^{(z)}$ as the intersection of the red and blue contour events corresponding to the individual components of the possible QoI maps. It is visually evident from the top-left plot of Figure 4 that the pair of QoI that minimizes $\overline{M_{Q(z)}}$ also produces the inverse image with the sharpest corners, i.e., it has the highest skewness. In the top-middle plot of Figure 4 we observe the opposite effect, the pair of QoI that minimizes the skewness $\overline{S_{Q(z)}}$ maximizes $\overline{M_{Q(z)}}$. In the top-right plot of Figure 4, we observe a pair of QoI that produces reasonably low $\overline{S_{Q(z)}}$ and $\overline{M_{Q(z)}}$ simultaneously.

4.2.2 Nonlinear

Let $Q : \mathbb{R}^2 \rightarrow \mathbb{R}^{10}$ be *nonlinear* where each component is a polynomial of the form

$$Q_j = r_{j,0}\lambda_1^5 + r_{j,1}\lambda_2^3 + r_{j,2}\lambda_1^3\lambda_2 + r_{j,3}\lambda_1 + r_{j,4}\lambda_2 + r_{j,5} \text{ for } j = 0, 1, \dots, 9,$$

where the coefficients $r_{j,k}$ are fixed random numbers in $[-1, 1]$. Let $B^{(z)} \in \mathcal{B}_{\mathcal{D}^{(z)}}$ be the rectangle of uncertainty in the data space defined by the z^{th} possible pair of QoI. Again, we assume the uncertainty in each QoI is the same so that $\mu(B^{(z)})$ is constant for all z . Now that Q is nonlinear, we approximate both the local Jacobians $J_{\lambda^{(i)}}$ and the integrals in Eqs. (14) and (26). We take 1000 uniform random samples $\lambda^{(i)}$ in Λ and then compute the approximate Jacobian at $N = 100$ of these samples using a Radial Basis Function (RBF) interpolation method. Specifically, for a given $\lambda^{(i)} \in \Lambda$, we use the 20 nearest neighbors of $\lambda^{(i)}$ (from the original 1000 samples) to approximate the gradient for each QoI and subsequently construct the Jacobian.

Remark 4. *Because Q is a vector valued polynomial function, we could analytically determine the exact gradient vectors for each component and simply evaluate at each $\lambda^{(i)}$. However, the RBF method (or any finite difference method) is a more general approach that can be used to approximate gradient vectors for maps Q defined by functionals of solutions to partial differential equations where we cannot analytically determine the gradients of each component.*

We illustrate the various parts of the optimization problem by separately solving four optimization problems; minimize $\overline{M_{Q(z),N}}$, minimize $\overline{S_{Q(z),N}}$, minimize Equation 30 with $\omega = 0.5$, and maximize Equation 30 with $\omega = 0.5$. The results are summarized in Figure 5 and Table 3. In the bottom left of Figure 5 we see the location in Y_ω of $s_{5,9} = (\overline{S_{Q_{5,9},N}}, \overline{M_{Q_{5,9},N}})$. Although this pair of QoI minimizes $\overline{M_{Q(z),N}}$ it is obvious from the scatter plot there are pairs of QoI that have similar average μ_Λ -measure that have

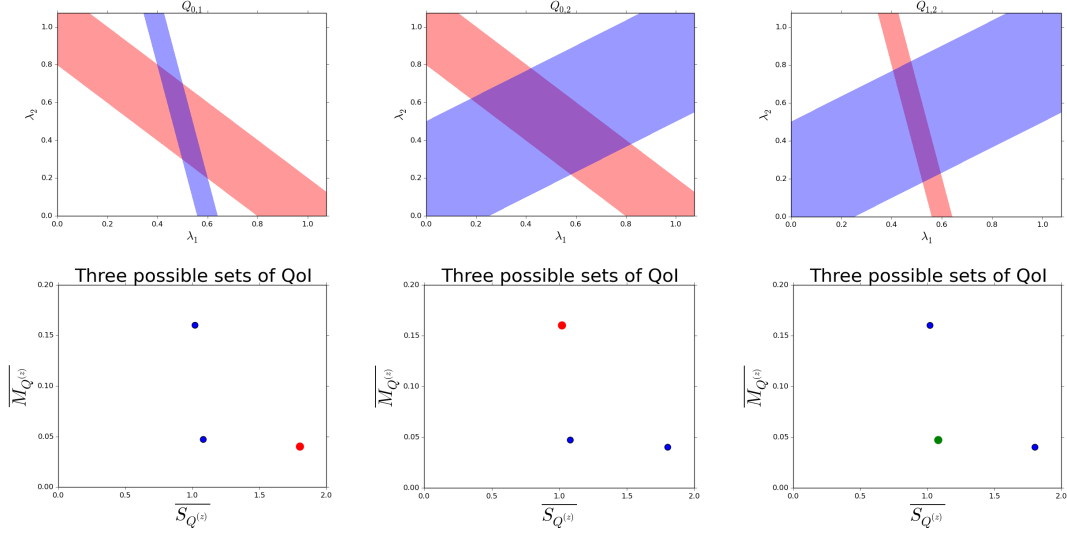


Figure 4: On the top we show the inverse density for the three possible pairs of QoI. On the bottom we show the point in Y_ω defined by each pair of QoI. (left): The pair of QoI that minimizes $\overline{M_{Q^{(z)}}}$ ($Q_{0,1}$). (middle): The pair of QoI that minimizes $\overline{S_{Q^{(z)}}}$ ($Q_{0,2}$). (right): The pair of QoI that minimizes the distance defined in Eq.(30) with $\omega = 0.5$ ($Q_{1,2}$).

Pair	$\overline{M_{Q^{(z)}}}$	$\overline{S_{Q^{(z)}}}$	$d_{Y_\omega}(p, y_z)$
$Q_{0,1}$	$4.0E - 2$	$1.80E + 0$	$4.84E - 1$
$Q_{0,2}$	$1.6E - 2$	$1.02E + 0$	$1.57E - 1$
$Q_{1,2}$	$4.7E - 2$	$1.08E + 0$	$1.20E - 1$

Table 2: $\overline{M_{Q^{(z)}}}$, $\overline{S_{Q^{(z)}}}$, and $d_{Y_\omega}(p, y_z)$ for each of the three possible pairs of QoI.

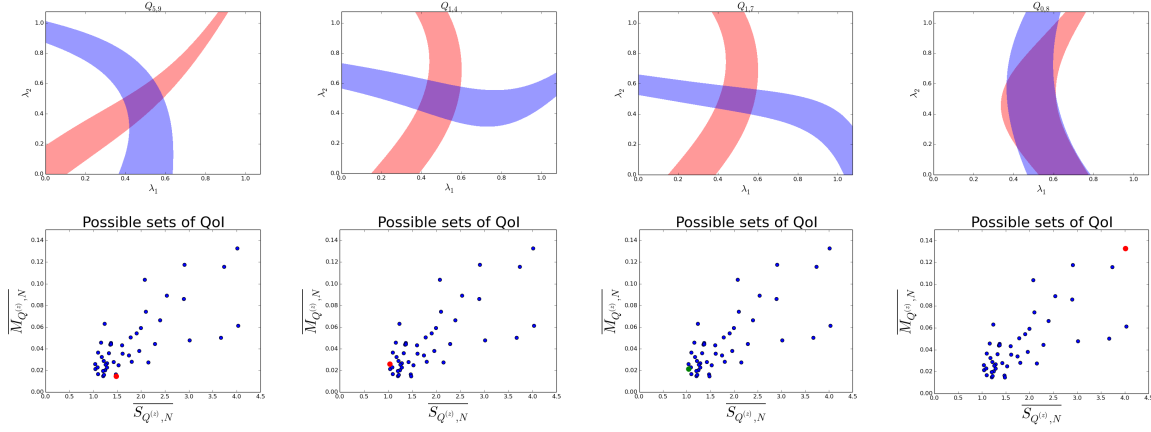


Figure 5: *On the top we show the inverse image defined by four possible pairs of QoI. On the bottom we show the point in Y_ω defined by each pair of QoI. (left): The pair of QoI that minimizes $\overline{M_{Q^{(z)},N}}$ ($Q_{5,9}$). (middle-left): The pair of QoI that minimizes $\overline{S_{Q^{(z)},N}}$ ($Q_{1,4}$). (middle-right): The pair of QoI that minimizes the distance defined in Eq.(30) with $\omega = 0.5$ ($Q_{1,7}$). (right): The pair of QoI that maximizes the distance defined in Eq.(30) with $\omega = 0.5$ ($Q_{0,8}$).*

Pair	$\overline{M_{Q^{(z)},N}}$	$\overline{S_{Q^{(z)},N}}$	$d_{Y_\omega}(p, y_z)$
$Q_{5,9}$	$1.50E - 2$	$1.482E + 0$	$3.40E - 1$
$Q_{1,4}$	$2.60E - 2$	$1.040E + 0$	$6.40E - 2$
$Q_{1,7}$	$2.10E - 2$	$1.044E + 0$	$6.30E - 2$
$Q_{0,8}$	$1.33E - 1$	$4.016E + 0$	$8.68E - 1$

Table 3: $\overline{M_{Q^{(z)},N}}$, $\overline{S_{Q^{(z)},N}}$, and $d_{Y_\omega}(p, y_z)$ for the four pairs of QoI considered in Figure 5.

much smaller average skewness. In the bottom right we see the location in Y_ω of $s_{0,8}$ that corresponds to the pair of QoI that maximizes the sum of the average μ_Λ -measure and the average skewness. The corresponding inverse image, seen in the top right of Figure 5, clearly has both the largest measure of support and the largest skewness of the four pairs considered.

5 Maps Defined by the Solution of a Physics-Based Model

In this section we consider the inverse density of sets under functions defined by the solutions to partial differential equations. We first consider a simple time dependent diffusion model with uncertain diffusion coefficient that allows us to understand the results intuitively. Then we consider the ADvanced CIRCulation model for oceanic, coastal, and estuarine waters (ADCIRC) which incorporates a spatially varying bottom friction model. The ADCIRC model depends on many parameter fields, we focus on the bottom friction parameter due to its inherent uncertainty.

5.1 Time Dependent Diffusion

5.1.1 The Model

We consider the time dependent diffusion equation

$$\rho c \frac{\partial T}{\partial t} = \nabla \cdot (\kappa \nabla T) + f, \quad x \in \Omega, t \in (t_0, t_f)$$

$$f(x) = \begin{cases} Ae^{-\frac{(x_0 - p_0)^2 + (x_1 - p_1)^2}{w}} & \text{if } t \leq t_{source} \\ 0 & \text{if } t > t_{source} \end{cases}$$

$$T(x; 0) = 0$$

$$\frac{\partial T}{\partial n} = 0, x \in \partial\Omega$$

where $\Omega = [-\frac{1}{2}, \frac{1}{2}] \times [-\frac{1}{2}, \frac{1}{2}]$ represents a thin plate made of some alloys, ρ is the density of the plate, c is the heat capacity, and κ is the thermal conductivity. The forcing function f represents an external source with the following source parameters: A is the amplitude, \bar{p} is the position, and w is the width. The external source is positioned at the center of the plate. We remove the external source at time $t_{source} = \frac{t_f}{2}$. The homogeneous Neumann boundary conditions model perfect insulation around the boundary.

We assume that the square plates are manufactured by welding together two rectangular plates of the same alloy. However, due to imperfections in the original manufacturing

process, we assume that the alloy composition varies significantly from plate to plate. Thus, we let the thermal conductivity κ vary in space as described below. We assume that the left-half of Ω contains an alloy with thermal conductivity κ_0 and the right-half of Ω contains an alloy with thermal conductivity κ_1 on the right, both of which are uncertain within $[0.01, 0.2]$. Thus, the model parameters for this problem are κ_0 and κ_1 and we have defined our parameter space, $\Lambda = [0.01, 0.2] \times [0.01, 0.2] \subset \mathbb{R}^2$. See Figure 6 where the green and purple colors indicate the two rectangular plates that are welded together to form the single plate defining the physical domain Ω

5.1.2 The Inverse Problem

Our goal is to determine which pair of QoI produce the best solution to the inverse problem, i.e., determine which two points in space time we should record temperature measurements in order to obtain the “best” inferences about the thermal conductivities of each side of the plate. Intuitively, we expect to take one temperature measurement on each side of the plate, but it is not clear at what time steps each temperature measurement should be recorded or exactly where we should place the sensors on each side of the plate. We limit ourselves to a discrete set of possible temperature measurements in space-time. We choose 20 points on the plate in space, see Figure 6, and gather these temperature measurements at each of 20 time steps. Thus, \mathcal{Q} is defined by 400 possible QoI values in space-time for which it is only possible to record two.

To make inferences about the optimal experimental design we must solve the model for a set of samples in the parameter space. For the numerical solution of the model, we use the state of the art open source finite element software FEniCS [1] with a 40×40 triangular mesh, piecewise linear finite elements and a Crank Nicolson time stepping scheme. We take 5000 uniform random samples from Λ and run the simulation for each sample. We mathematically model the sensors recording data at a point in space-time as a functional of the solution by computing the average temperature on a small disc about a point in space but at a fixed time, i.e., each QoI is modeled as

$$Q_{(p_i; t_i)} = \frac{1}{\mu_{\Omega}(\mathbf{B}_r(p_i))} \int_{\Omega} T(x; t_i) \chi_{\mathbf{B}_r(p_i)} dx.$$

This yields a map $Q : \Lambda \rightarrow \mathcal{D}$ where Λ is 2-dimensional and \mathcal{D} is 400-dimensional. Thus, we have $\binom{400}{2} = 79,800$ possible pairs of QoI to consider.

We approximate $\overline{M_{Q(z)}}$ and $\overline{S_{Q(z)}}$, for each z , using the $N = 5000$ samples for which we solved the model. We solve three optimization problems; minimize $\overline{M_{Q(z), N}}$, minimize $\overline{S_{Q(z), N}}$, and minimize the distance defined in Eq. (30) with $\omega = 0.5$. Recall these minimization problems tell us *on average* the best possible set of QoI to use in the inverse problem before any physical experiment occurs.

The results of the optimization problems are shown in Table 4. As expected, each pair of QoI found has one point on the left side of the plate and one point on the right. However, there are notable differences in space-time locations of each optimal

QoI Locations

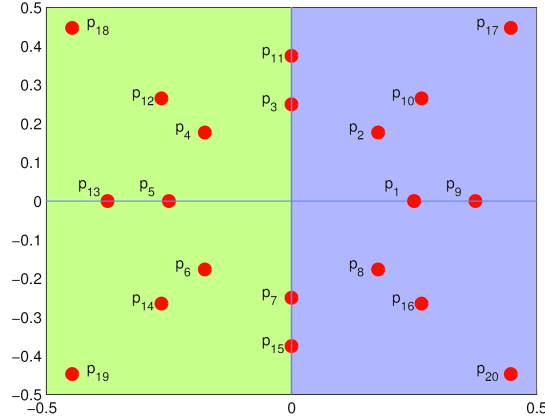


Figure 6: *The QoI locations in the domain Ω .*

pair of QoI. The pair of QoI that minimizes $\overline{M_{Q^{(z)},N}}$ is produced by measurements made near the end of the simulation. The pair of QoI that minimizes $\overline{S_{Q^{(z)},N}}$ is produced by measurements made at the first time step. Intuitively, this makes sense because at t_1 the temperature measurements are primarily determined by local alloy properties, i.e., not enough time has passed for the left side of the plate to influence the temperature measurements on the right side of the plate and vice versa.

We solve the inverse problem three times, once for each pair of QoI described above. We simulate real data, i.e., a temperature measurement made on a given plate with uncertain thermal conductivities in a laboratory experiment, by mapping a random point $\lambda_{ref} \in \Lambda$ forward to $Q^{(z)}(\lambda_{ref}) = q_{ref}^{(z)} \in \mathcal{D}^{(z)}$. We assume the uncertainty in this *observed* datum is the same for each QoI and let this uncertainty be 0.5 degrees, i.e., the 2-dimensional square we invert for each pair of QoI is centered at $q_{ref}^{(z)}$ and has side length of 0.5. We solve the inverse problem using the open source BET [6] python package for each of the resulting pairs of QoI, see Figure 7.

In the top row we see in purple the inverse image for each pair of QoI approximated with $N = 5000$ uniform samples. Notice in the top middle, the pair that minimizes the average skewness, the inverse image is the entire parameter space, i.e., we have not reduced the uncertainty in our model input parameters at all. Although this inverse image can be well approximated by relatively few samples, it is useless as it gives no new information about the location of the parameters that produced the observed datum. In the bottom row we show the location of $y_z \in Y_\omega$ for each pair of QoI. In the bottom right, notice the apparent smooth curve being defined by these points in Y_ω . In multicriteria optimization this is referred to as the Pareto frontier, and we see in all three cases our optimal choice of pair of QoI produces a point in Y_ω that lies very near the Pareto frontier.

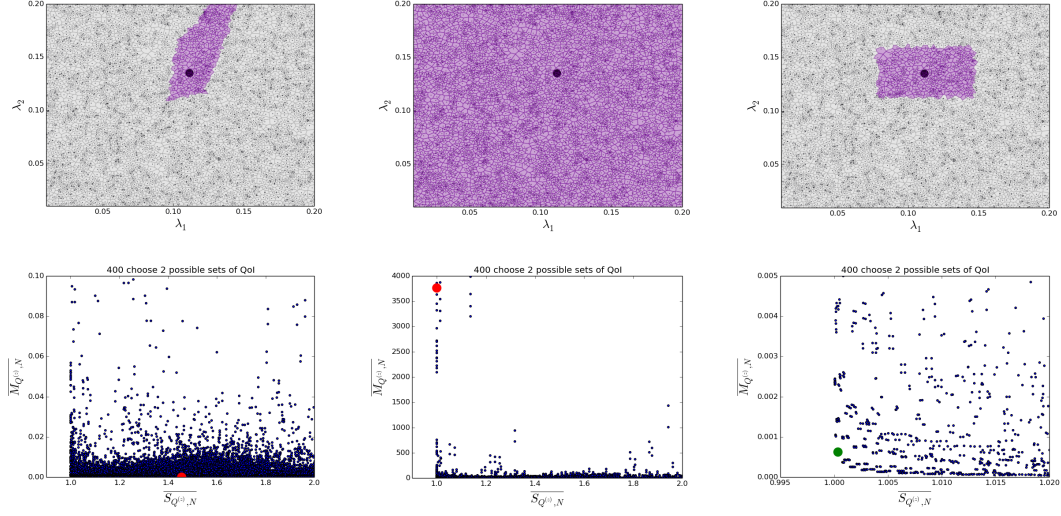


Figure 7: (top): The inverse image approximated using the BET package. (bottom): The space Y_ω defined in Section 4. (left): The pair of QoI that minimizes $\overline{M_{Q^{(z)},N}}$. (middle): The pair of QoI that minimizes $\overline{S_{Q^{(z)},N}}$. (right): The pair of QoI that minimizes the distance defined in Eq.(30) with $\omega = 0.5$.

QoI Locations	$\overline{M_{Q^{(z)},N}}$	$\overline{S_{Q^{(z)},N}}$	$d_{Y_\omega}(p, y_z)$
$Q_{(p_{18},t_{14})}, Q_{(p_{17},t_{15})}$	$1.72E - 05$	$1.46E + 00$	$3.31E - 01$
$Q_{(p_1,t_1)}, Q_{(p_{12},t_1)}$	$3.76E + 03$	$1.00E + 00$	$9.99E - 01$
$Q_{(p_{17},t_4)}, Q_{(p_{18},t_4)}$	$6.26E - 04$	$1.0004E + 00$	$9.84E - 04$

Table 4: $\overline{M_{Q^{(z)},N}}$, $\overline{S_{Q^{(z)},N}}$, and $d_{Y_\omega}(p, y_z)$ for the three pairs of QoI considered in Figure 7.

QoI Locations	$\frac{\mu_\Lambda(\text{supp}((Q^{(z)})^{-1}(B)))}{\mu_\Lambda(\Lambda)}$	Prediction Interval
$Q_{(p_{18}, t_{14})}, Q_{(p_{17}, t_{15})}$	$6.68E - 2$	$[18.3, 24.0]$
$Q_{(p_1, t_1)}, Q_{(p_{12}, t_1)}$	$1.00E + 0$	$[0.06, 25.3]$
$Q_{(p_{17}, t_4)}, Q_{(p_{18}, t_4)}$	$8.62E - 2$	$[16.9, 23.1]$

Table 5: *Predictions made by propagating forward each of the three inverse images in Figure 7.*

5.1.3 The Prediction Problem

As mentioned in Section 3, following solution to the stochastic inverse problem, we generally want to make, and quantify uncertainty in, predictions that can be formulated as solutions to a stochastic forward problem. Typically these predictions correspond to unobservable QoI, e.g., the maximum storm surge elevation for a future hurricane being forced by currently unknown winds and tides. For simplicity, we consider a prediction problem for a similar QoI in the heating of a thin plate as described above. We denote the prediction QoI as $Q^{(p)}$ which represents the average temperature along the right side of the plate at the final time step with the external source now located at the bottom left of the plate.

$$Q^{(p)} = \frac{1}{\mu_\Omega(E)} \int_\Omega T(x; t_{20}) \chi_E dx,$$

$$f(x) = \begin{cases} Ae^{-\frac{(x_0 + \frac{1}{2})^2 + (x_1 + \frac{1}{2})^2}{w}} & \text{if } t \leq t_{source} \\ 0 & \text{if } t > t_{source}. \end{cases}$$

Here, E is a small vertical strip in Ω along the right boundary. For each inverse solution found in Section 5.1.2 we have a corresponding set of samples from the original $N = 5000$ that lie inside the inverse density. For each of these samples, we run the simulation and evaluate $Q^{(p)}$. The range of the results for each of the three inverse images is shown in Table 5. In the first column we give the pair of QoI used to solve the stochastic inverse problem. In the second column we show the relative measure of the support of the probability measure solving the stochastic inverse problem. In the third column we show the interval of possible prediction values for $Q^{(p)}$. We see the pair of QoI that minimizes the sum of the average μ_Λ -measure and the average skewness (last row) produces a prediction interval about 5 times smaller than propagating the entire parameter space forward (second row). Also, this pair of QoI produces a prediction interval of similar precision (as determined by length of the interval) to the prediction interval corresponding to the first row, which we would expect to be the best but perhaps the most computationally complex to use.

5.2 Coastal Ocean Model

5.2.1 The Model

Coastal ocean models numerically solve the shallow water equations (SWEs), which model the flow of water processes on domains with vertical length scales that are negligible relative to the horizontal length scales. Integrating out the depth results in a first-order hyperbolic continuity equation coupled to the momentum equations for horizontal depth-averaged velocities given by

$$\frac{\partial H}{\partial t} + \frac{\partial}{\partial x}(Q_x) + \frac{\partial}{\partial y}(Q_y) = 0, \quad (33)$$

$$\begin{aligned} \frac{\partial Q_y}{\partial t} + \frac{\partial U Q_x}{\partial x} + \frac{\partial V Q_x}{\partial y} - f Q_x = \\ -gH \frac{\partial(\zeta + P_s/g\rho_0 - \alpha\eta)}{\partial x} + \frac{\tau_{sx}}{\rho_0} - \frac{\tau_{bx}}{\rho_0} + M_x - D_x - B_x, \end{aligned} \quad (34)$$

$$\begin{aligned} \frac{\partial Q_x}{\partial t} + \frac{\partial U Q_y}{\partial x} + \frac{\partial V Q_y}{\partial x} - f Q_x = \\ -gH \frac{\partial(\zeta + P_s/g\rho_0 - \alpha\eta)}{\partial y} + \frac{\tau_{sy}}{\rho_0} - \frac{\tau_{by}}{\rho_0} + M_y - D_y - B_y. \end{aligned} \quad (35)$$

Here, ζ is the free surface departure from the geoid, h is the bottom surface departure from the geoid, $H = \zeta + h$ is the total water column height, U_i is the depth-averaged velocity in the x_i direction, f is the Coriolis parameter, P_s is the atmospheric pressure at the free surface, ρ_0 is the reference density of water, α is the effective earth elasticity factor, η is the Newtonian equilibrium tide potential, τ_{sx_i} are the imposed surface stresses, τ_{bx_i} are the bottom stress components, M_{x_i} is the vertically-integrated lateral stress gradient, D_{x_i} is the momentum dispersion, and B_{x_i} is the vertically-integrated baroclinic pressure gradient.

In the ADvanced CIRCulation (ADCIRC) model the continuity equation, Equation 33, is replaced by the second order hyperbolic generalized wave continuity equation (GWCE), Eq. (36) [3]. Together, the GWCE and the momentum equations, Eqs. (34) and (35), define the modified form of the SWEs solved by ADCIRC. The GWCE and momentum equations are discretized in space by a piecewise linear triangular mesh. The time stepping is done with centered finite differences for the GWCE and forward finite difference for the momentum equations.

Parameter	Range of Manning's n values	Land Classification
λ_1	[0.0396, 0.21]	Low-intensity developed
λ_2	[0.0594, 0.315]	Evergreen forest
λ_3	[0.0495, 0.2625]	Palustrine forested wetland
λ_4	[0.00825, 0.04375]	Open water

Table 6: *Manning's n coefficient ranges for the subdomain.*

$$\begin{aligned}
& \frac{\partial^2 \zeta}{\partial t^2} + \tau_0 \frac{\partial \zeta}{\partial t} + \frac{\partial}{\partial x} \left(-\frac{\partial U Q_x}{\partial x} - \frac{\partial V Q_x}{\partial x} + f Q_y - g H \frac{\partial(\zeta + P_s/g\rho_0 - \alpha\eta)}{\partial x} \right. \\
& \left. + \frac{\tau_{sx}}{\rho_0} - \frac{\tau_{bx}}{\rho_0} + M_x - D_x - B_x + \tau_0 Q_x \right) + \frac{\partial}{\partial y} \left(-\frac{\partial U Q_y}{\partial x} - \frac{\partial V Q_y}{\partial x} + f Q_x \right. \\
& \left. - g H \frac{\partial(\zeta + P_s/g\rho_0 - \alpha\eta)}{\partial y} + \frac{\tau_{sy}}{\rho_0} - \frac{\tau_{by}}{\rho_0} + M_y - D_y - B_y + \tau_0 Q_y \right) \\
& - U H \frac{\partial \tau_0}{\partial x} - V H \frac{\partial \tau_0}{\partial y} = 0.
\end{aligned} \tag{36}$$

A driving application of the ADCIRC model is predicting maximum storm surge elevations during extreme weather events such as hurricanes. Accurate estimation of storm surge elevations requires accurate estimation of model parameters. The Manning's n coefficients of roughness are of particular importance. They enter into the SWEs through the bottom stress components τ_{bxi} in the momentum equations. The Manning's n coefficient is a highly variable spatial parameter dependent on the surface characteristics of the seabed and is inherently uncertain.

For the purposes of this work, the derivation of how the Manning's n coefficient enters into the momentum equation is not particularly important. In the end, as we change the Manning's n coefficient we see changes in the maximum storm surge elevations in the physical domain.

5.2.2 The Inverse Problem

We consider a subdomain of the Gulf Coast that has four dominant land classifications and therefore four dominant Manning's n coefficients that enter into the bottom stress components in the momentum equations, see Table 6. We run ADCIRC to simulate the storm surge under the forcing of Hurricane Gustav and record maximum storm surge elevations at 194 spatial locations, see Figure 8. Our goal is to determine which set of four spatial locations provides us with data that produces the best solution to the inverse problem.

Remark 5. *The results from this section are from Lindley Graham's PhD Thesis [10] and were computed during the summer of 2015. The algorithm used to determine the optimal set of QoI has since developed into the methods described in Sections 3 and*

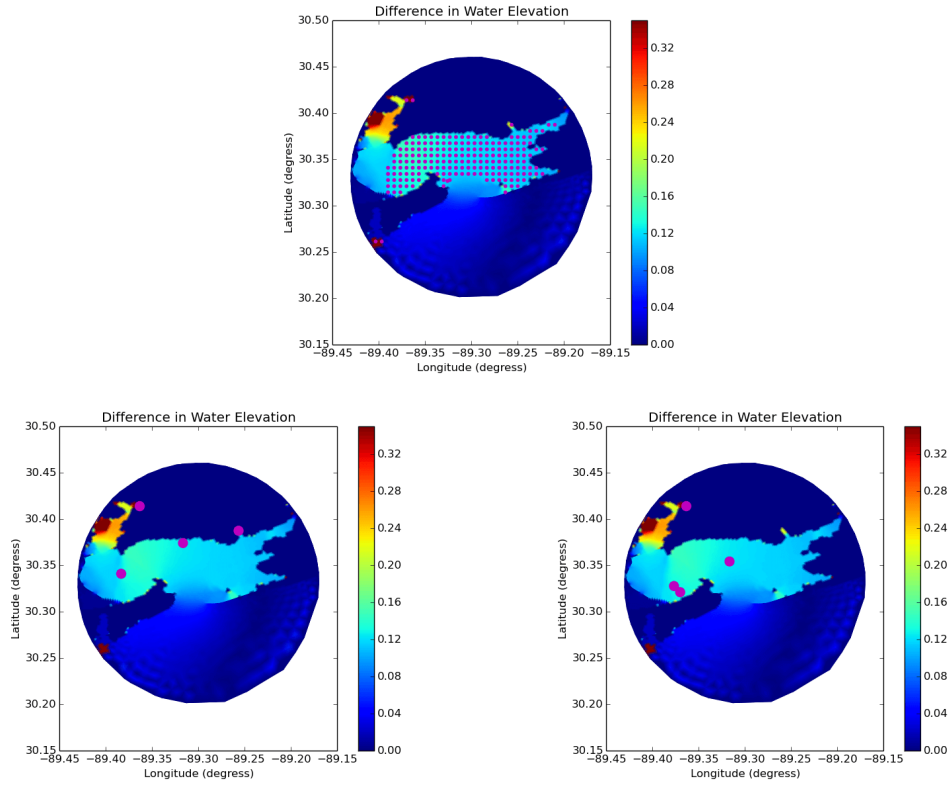


Figure 8: (top): 194 potential station locations. (bottom left): Stations [11, 40, 160, 91] define the optimal observation network. (bottom right): Stations [16, 40, 22, 88] define a suboptimal observation network.

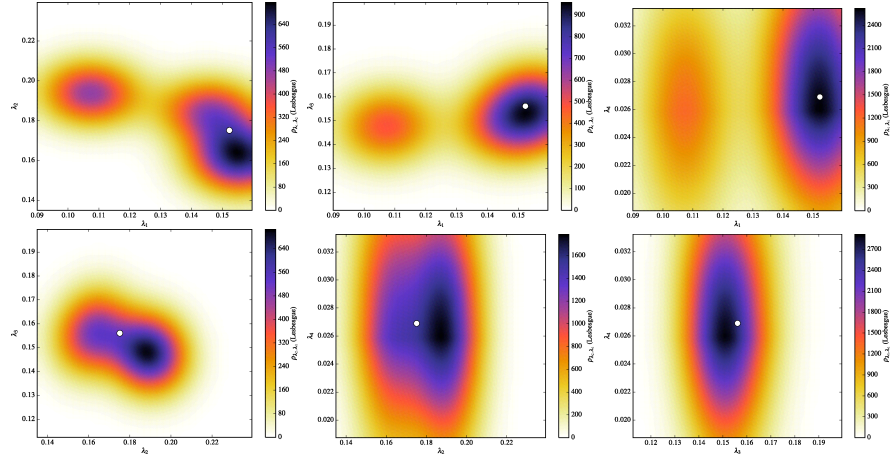


Figure 9: *Plots of the marginals of the inverse solution using uniform samples for optimal stations [11, 40, 160, 191]. Here, ρ_D is defined as a uniform density on a small rectangular box centered at the reference QoI values associated with $\lambda_{\text{ref}} = (0.1523, 0.1751, 0.1561, 0.0269)$. The reference value is illustrated by a white circle. (left): In order from top to bottom $(\lambda_1, \lambda_2), (\lambda_1, \lambda_4), (\lambda_2, \lambda_4)$. (right): In order from top to bottom $(\lambda_1, \lambda_3), (\lambda_2, \lambda_3), (\lambda_3, \lambda_4)$.*

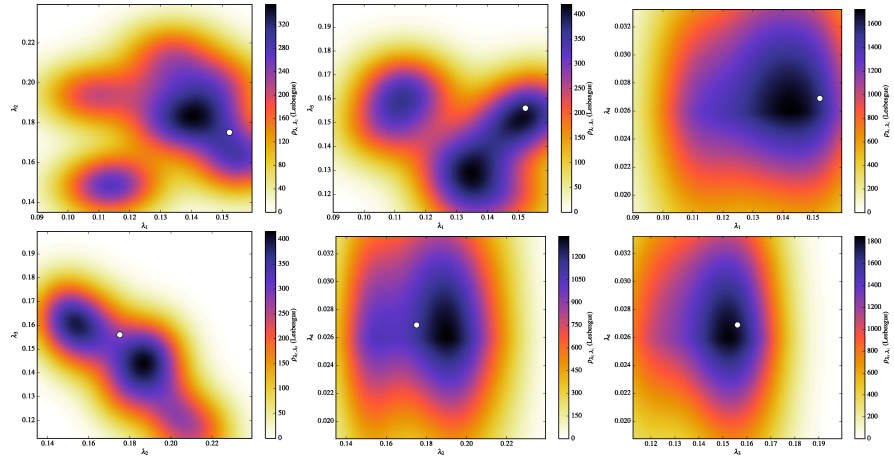


Figure 10: *Plots of the marginals of the inverse solution using uniform samples for suboptimal stations [16, 40, 22, 88]. Here, ρ_D is defined as a uniform density on a small rectangular box centered at the reference QoI values associated with $\lambda_{\text{ref}} = (0.1523, 0.1751, 0.1561, 0.0269)$. The reference value is illustrated by a white circle. (left): In order from top to bottom $(\lambda_1, \lambda_2), (\lambda_1, \lambda_4), (\lambda_2, \lambda_4)$. (right): In order from top to bottom $(\lambda_1, \lambda_3), (\lambda_2, \lambda_3), (\lambda_3, \lambda_4)$.*

4. *Although these results were found using a less tuned algorithm, they still display fundamental characteristics of inverse solutions defined by optimal and suboptimal sets of QoI.*

In Figure 8 we see three plots of the subdomain. The points in each plot represent particular stations of interest, the smooth coloring of the domain represents the difference in the computationally approximated maximum storm surge elevation over 480 uniform samples in Λ . Notice in the top image we do not place potential station locations in regions of the subdomain that do not display significant sensitivity in storm surge elevations as the Manning's n coefficients (parameters) vary.

In Figure 9 we see the marginal densities computed from the inverse solution for the optimal set of QoI seen in the bottom left of Figure 8. In Figure 10 we see the marginal densities computed from the inverse solution for the suboptimal set of QoIs seen in the bottom right of Figure 8. To clearly draw distinctions between these separate solutions, we examine the top left plots in Figures 9 and 10. Note that although both marginal densities contain the reference point λ_{ref} within their supports, the solution from the optimal QoI does a much better job of concentrating high probability in a smaller region of Λ containing this reference point. The relative measure of the support of the density computed for the optimal set of stations is $\mu_{\Lambda}((Q^{(opt)})^{-1}(B))/\mu_{\Lambda}(\Lambda) = 7.904E-3$, where as for the suboptimal set $\mu_{\Lambda}((Q^{(subopt)})^{-1}(B))/\mu_{\Lambda}(\Lambda) = 1.917E-2$. Using the optimal set of stations has reduced the relative measure of the support of the inverse density by approximately a factor of 3.

6 Conclusion

In [7] the local skewness of the inverse image of a generalized rectangle in some \mathcal{D} was defined. We have extended this initial work to quantify the global effect of skewness which is a way of quantifying the accuracy we can obtain in solution to the stochastic inverse problem from a finite number of samples. We also developed a way to quantify the precision in solutions to stochastic inverse problems in terms of measures of supports of densities and/or high probability events. A multicriteria optimization problem was defined to determine the best QoI map from the space of theoretically possible QoI that simultaneously reduces both the skewness and measure of the support of the inverse density. Several numerical examples demonstrated the effect of the QoI choice on the solution to the stochastic inverse problem.

Future Work

As mentioned in a previous remark, the choice of weights used in Eq. (30) is a current topic of discussion. It is clear as more computational resources are available we desire to reduce the μ_{Λ} -measure of the inverse image more than the skewness, i.e., we set $\omega \ll 1$. A fundamental problem is to determine if the value of ω can be solved for directly as a function of the number of model solves available.

In this work we considered the discrete optimization problem, i.e., given m sensors and d theoretical QoI to choose from, determine the optimal set to improve solutions to the corresponding stochastic inverse problem. The continuous analogue is simply choose m optimal QoI from infinitely many possibilities. As a concrete example, suppose that for the PDE in Section 5.1 we can place two temperature sensors at *any* point in space-time, and we wish to choose the best two points in space-time to record data. This problem is approached by considering the function $R : \mathcal{Q} \rightarrow Y_\omega$, where \mathcal{Q} is now the set of infinitely many possible sets of QoI, and finding local minima of this function. In the case of QoI being represented by some specific type of functional at a point in space time (temperature measurement), the space \mathcal{Q} has dimension $(\dim(\Omega) + 1) \times m$ where Ω is the spatial domain of the model and m is the number of QoI to be chosen. Although the dimension space \mathcal{Q} increases quickly as m increase, there appears to be symmetry to exploit in this space.

In high dimensional parameter spaces any set implicitly defined by the solution to a stochastic inverse problem, skewed or not, is difficult to approximate with a reasonably low number of samples. In [8] adaptive sampling algorithms are used to combat this curse of dimensionality. With the ability to place samples near the boundary of sets in Λ the error in the approximation of these sets is greatly reduced. With the addition of gradient information the efficiency of these adaptive sampling algorithms is improved. Preliminary results indicate the efficiency is improved further as the skewness of the set is reduced. A thorough numerical exploration of the effects of skewness on the efficiency of adaptive sampling algorithms is an obvious path to explore.

In many driving applications the collection of field data is difficult, time consuming, and expensive. Obtaining a good solution to the inverse problem while using as little data as possible is greatly desired. In the examples in Section 5 we look to determine the best set of m QoI to use to solve the inverse problem, however, possibly a set of $m - 1$ QoI produces a very similar inverse solution, or a set of $m - 2$. This information can greatly reduce the resources needed to gather the field data required. With the $M_{Q^{(z)}}$ and $S_{Q^{(z)}}$ defined as in Section 3, can we use $M_{Q^{(z)},k}$ and $S_{Q^{(z)},k}$ (the μ_Λ -measure and skewness of the map $Q^{(z)}$ without the k^{th} QoI) to draw conclusions about the maximum number of useful QoI to use from a given set?

References

- [1] Martin S. Alnæs, Jan Blechta, Johan Hake, August Johansson, Benjamin Kehlet, Anders Logg, Chris Richardson, Johannes Ring, Marie E. Rognes, and Garth N. Wells. The fenics project version 1.5. *Archive of Numerical Software*, 3(100), 2015.
- [2] J. Breidt, T. Butler, and D. Estep. A measure-theoretic computational method for inverse sensitivity problems I: Method and analysis. *SIAM Journal on Numerical Analysis*, 49:1836–1859, 2011.
- [3] T. Butler, M.U. Altaf, C. Dawson, I. Hoteit, X. Luo, and T. Mayo. Data assimilation within the advanced circulation (ADCIRC) modeling framework for hurricane storm surge forecasting. *Monthly Weather Review*, 140:2215–2231, 2012.
- [4] T. Butler, D. Estep, and J. Sandelin. A computational measure-theoretic approach to inverse sensitivity problems II: A posteriori error analysis. *SIAM Journal on Numerical Analysis*, 50:22–45, 2012.
- [5] T. Butler, D. Estep, S. Tavener, C. Dawson, and J.J. Westerink. A measure-theoretic computational method for inverse sensitivity problems III: Multiple output quantities of interest. *SIAM/ASA Journal on Uncertainty Quantification*, 2:174–202, 2014.
- [6] T. Butler, D. Estep, S. Tavener, T. Wildey, C. Dawson, and L. Graham. Solving stochastic inverse problems using sigma-algebras on contour maps. 2015. arXiv preprint arXiv:1407.3851.
- [7] T. Butler, L. Graham, D. Estep, C. Dawson, and J.J. Westerink. Definition and solution of a stochastic inverse problem for the manning’s n parameter field in hydrodynamic models. *Advances in Water Resources*, 78:60–79, 2015.
- [8] T. Butler, L. Graham, S. Mattis, and S. Walsh. A measure-theoretic interpretation of sample based numerical integration with applications to inverse and prediction problems under uncertainty. In preparation.
- [9] C. Dellacherie and P.A. Meyer. *Probabilities and Potential*. North-Holland Publishing Co., Amsterdam, 1978.
- [10] L. Graham. *Adaptive Measure-Theoretic Parameter Estimation for Coastal Ocean Modeling*. PhD thesis, Institute for Computational Engineering and Sciences, The University of Texas at Austin, Austin, TX 78712, 2015.